(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 887 958 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
22.01.2003 Bulletin 2003/04

(51) Int Cl.[7]: **H04H 9/00**

(21) Application number: **98810563.1**

(22) Date of filing: **19.06.1998**

(54) **Method for the compression of recordings of ambient noise, method for the detection of program elements therein, devices and computer program therefor**

Verfahren für die Kompression der Aufnahmen von Umgebungsgeräuschen, Verfahren für die Erfassung von Programmelementen darin, Vorrichtung und Computer-Programm dafür

Méthode pour la compression des enregistrements du bruit ambiant et pour y détecter des éléments de programme, dispositif et logiciel pour la mise en oeuvre d'une telle méthode

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU NL PT SE
Designated Extension States:
RO SI

(30) Priority: **23.06.1997 CH 152097**

(43) Date of publication of application:
30.12.1998 Bulletin 1998/53

(73) Proprietor: **LIECHTI AG**
CH-4566 Kriegstetten (CH)

(72) Inventor: **Bichsel, Martin**
8544 Bertschikon (CH)

(74) Representative:
AMMANN INGENIEURS-CONSEILS EN PROPRIETE INTELLECTUELLE SA BERNE
Schwarztorstrasse 31
3001 Bern (CH)

(56) References cited:

EP-A- 0 118 771	**WO-A-84/02793**
DE-A- 4 400 683	**FR-A- 2 715 016**
US-A- 3 919 479	**US-A- 4 450 531**

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The present invention refers to a method for the compression of an electric audio signal which is produced in the process of recording the ambient noise by means of an electroacoustic transducer, more particularly a microphone. Furthermore, the invention also refers to a device for carrying out the method, a computer program product for carrying out the method, and a data carrier containing such a computer program product.

[0002] In the field of audience research, which also comprises the acoustic perception of other media such as e.g. television, recordings of the acoustic environment of a panelist in a survey are used, i.e. the so-called hearing samples. The storage of these hearing samples on portable magnetic tape recorders is disclosed in US 5,023,929. The inconvenient of this method is that the tape recorder is relatively large although it is intended to be permanently carried by the participant.

[0003] Consequently, it would be preferable to integrate the hearing sample recorder or monitor in an appliance which is normally worn or is at least less visible. Such a possibility, namely the integration into a wristwatch, is mentioned in EP-A-0 598 682 to the applicant.

[0004] However, the mentioned application does not indicate how the hearing samples can be stored in the extremely narrow space and with the very limited energy available in a wristwatch or a similarly inconspicuous appliance over a considerable period of time such as at least a week. Although the specification mentions the need of compression procedures, known methods only are indicated.

[0005] US-4,450,531 discloses a method for comparing a broadcast signal with reference samples in order to determine the received program. Samples of the broadcast signal, captured e.g. by a tuner, a frequency band is filtered out and Fourier transformed. The reference samples are treated in almost the same way, however are further normalized to the power of each sample and only thereafter Fourier transformed. The two signal types are subjected to a correlation function, inversely Fourier transformed and the distances of the correlation peaks are determined. If they are equal to the length of the reference sample, the program samples and the reference samples contain the same program. It is further required to apply an additional test, i.e. to compare the power patterns obtained by taking the RMS power values of the two sample types at the correlation. This method merely intends to improve the correlation reliability, yet it does not take care of reducing the data volume, it even does not consider to store the samples. Finally, it requires a properly received signal. Detection of a broadcast program within environmental noise is not addressed.

[0006] DE-A-4,400,683 teaches a method, wherein amplitude-related values and frequency-weighing factors of hearing samples are calculated and stored for later evaluation. The use of two different value types, which are more or less indicative for the determination of the captured program in dependence on the more dynamic or static momentary characteristics of the program, requires at least additional efforts in the correlation process and entails increased power consumption for calculating two different series of values. It is therefore an object of the present invention to provide a method for the compression of hearing samples which in particular allows to obtain a high compression with minimal efforts with the safe recognition of program elements being essentially conserved.

[0007] This object is attained by a method according to claim 1. The further claims indicate preferred embodiments, devices for carrying out the method as in claim 24 a computer program product for carrying out the method and a data carrier as claimed in Claim 36 containing such a computer program product.

[0008] In the following, the same terminology as in EP-A-0 598 682 will be used. A hearing sample is basically a recording of the ambient noise e.g. by means of a microphone. In order to simplify the storage as well as the transmission to the evaluating center, however, it is preferred to have a succession of short recordings of the ambient noise or hearing samples which are recorded at certain times. Preferably, the recordings are effected at regular intervals of e. g. 1 minute, and have a constant duration of the order of, for example, 4 seconds, the information of the time of the recordings being stored together with the hearing sample.

[0009] According to the invention, the hearing samples are finally stored in an electronic memory in a digitized form. According to the invention, in order to reduce the amount of data to be stored, a normalization of the hearing samples in their original form or in a derived form (filtered, limited to selective frequency bands, digital or analog, etc.) to a predetermined range of values (e.g. amplitudes) D and a subsequent nonlinear transformation on a second range W is effected whose result, which is limited to the range W, is then stored in an electronic memory. The range W may be smaller or equal to D, but it is preferably substantially smaller.

[0010] Essentially, the non-linear transformation serves the purpose of amplifying sensitive areas of range D in such a manner that the more significant information provided by a signal whose value is comprised in such a sub-range of D is emphasized in the result, i.e. its resolution is increased.

[0011] Preferred further developments of the invention are as follows:

A: The nonlinear mapping is characterized by a decreasing slope dW/dD for increasing values in D, e.g. similar to the logarithmic function. Essentially, the range of small values in D is thereby mapped onto a relatively larger range in W and thus emphasized, whereas relatively large values in D are mapped on a relatively small range

in W only, i.e. their significance is attenuated.

B: The hearing samples are digitized immediately after recording (e.g. by a microphone) and analog processing (amplification; coarse filtering in preparation of the analog-digital conversion, etc.), resulting in a succession of numeric values. Each numeric value represents e.g. the momentary loudness of the ambient noise at a determined time.
Further processing is effected digitally by digital circuits, program controlled processors, or combinations thereof.

C: The amplitude or loudness values are transformed into energy values e.g. by squaring. The energy values are submitted to a low pass filtering and subsequently differentiated, the differentiation preferably being simulated by a difference calculus. The resulting energy variation values indicate the variation of the low-frequency proportion of the energy content in time.

D: The group of the energy variation values of a hearing sample, or only a part thereof, is normalized with respect to the maximum value of the values within the (partial) group. For this purpose, the maximum value is determined and all values of the group are divided by this maximum value. Simultaneously, the normalized values are mapped on a given range of numbers corresponding to the range D, e.g. the numbers between -128 and +127, so that the following arithmetic operations involve only integers. The number of values in these numerical ranges D is therefore preferably equal to powers of 2 (in the example: $256 = 2^8$ values) which are particularly advantageous in the case of binary digital processing. In order to perform this combination of normalizing and of imaging, the values of a group are multiplied by a factor which results from the division of the limit of the numeric range (i.e. 128 in the example) by the maximum value within the group.

E: The results of this step are again mapped on a further, smaller range of values W, e.g. the numerical range from 0 to 15 comprising $2^4 = 16$ numbers. On account of the fixed and relatively small number of values of the input data of this step, a so-called look-up table may be used for this second mapping.
Overall, it follows from the preceding that each numerical value of the hearing samples is reduced to a relatively short binary number (of 4 bits in the example).

F: Further optimizations are applied, such as e.g. taking the mean value of a plurality of values, only the mean value being further used. This also results in an important reduction of the number of values to be processed. On the digital level, such a filtering is simulated by a convolution.

G: Before or after being digitized at the input, the hearing sample is split into frequency bands or band signals. In a known manner, digital filterings may be effected by convolutions, and since the preferred convolutions represent low pass filterings, it is preferable to transmit less values to the following processing stages than are used for the convolution, preferably only one respective value.

[0012] The invention will be explained in more detail hereinafter by means of an exemplary embodiment and with reference to figures.

Fig. 1 shows a block diagram of a monitor according to the invention;

Fig. 2 shows the division into frequency bands;

Fig. 3 shows the conversion into energy values and the differentiation;

Fig. 4 shows the "normalizing quantization".

[0013] Fig. 1 shows a block diagram of a monitor 1. It may e.g. be intended to be integrated in a wristwatch, which is why monitor 1 comprises a clock circuit 2 which also serves as a time base for the signal processing, as well as a (liquid crystal) display 3. Commercially available components may be used for circuit 2 and display 3. A precise clock signal is generated by a quartz 4 in conjunction with an oscillator circuit which is integrated in clock circuit 2. Since a highly precise timing is required for the synchronization of the hearing samples to the comparative samples, a temperature compensation is provided in addition. The latter comprises a temperature sensor 5 which is connected to the clock circuit by means of an interface circuit 6. Interface circuit 6 essentially comprises an A/D converter.
[0014] Another important element for the monitor function is wearing detector 7. It may essentially consist of a sensor area on the wristwatch which detects the contact with the skin of the wearer. In the example, wearing sensor 7 is

connected to clock circuit 2 by means of an interface circuit 8, which implies that the clock circuit is capable of providing the time indications with an additional mark from the wearing sensor. It is also conceivable to directly connect the wearing sensor to the proper monitor circuit, e.g. to digital signal processor 9.

[0015] The clock signals which are required for the signal processing, in particular for signal processor 9, are derived from the time base clock, which is taken from a connection 10 of quartz 4, by a PLL (phase locked loop) circuit 11. The time and the date as well as the mark from the wearing sensor, as the case may be, are transmitted from clock circuit 2 to digital signal processor 9 by a serial data connection 12.

[0016] The hearing samples are stored in a flash memory. It is an important advantage with respect to the present application that flash memories are capable of storing data in a non-volatile manner and of deleting them again without the need of particular measures. A bus 14 allowing to transmit both data and addresses serves to connect flash memory 13 and signal processor 9.

[0017] A multiplexer 16 is connected by a second serial connection. Depending on the operational condition, the multiplexer connects signal processor 9 to the recording unit of the hearing samples or to interface circuit 17 by means of which the data exchange with the evaluating center is effected.

[0018] The recording unit consists of a microphone 18 and a following A/D converter unit 19 which in addition to the proper A/D converter may comprise amplifiers, filters (anti-aliasing filters) and other usual measures in order to ensure a digital signal which represents the recording by the microphone as correctly as possible.

[0019] Power supply 20 may be a battery (lithium cell) or the like. An accumulator in conjunction with a contactless charging system by means of electromagnetic induction or a photo cell is also conceivable.

[0020] To ensure the connection to the exterior, more particularly for the transmission of data to the evaluating center, monitor 1 is provided with a bidirectional data connection 21, a reset input 22, a synchronization input 23, and a power supply terminal 24. The presence of a power supply at terminal 24 is also used to make the monitor change to the data transmission mode. For example, the monitor may be connected to a base station which establishes a connection to an evaluating center e.g. by telephone. Another possibility consists in mailing the monitor to the center where it is connected to a reading station. On this occasion, besides the data transmission, a synchronization of clock circuit 2 to the clock of the center may be effected, as previously described in EP-A-0 598 682.

[0021] As shown in the illustration, the hearing sample processing unit including signal processor 9 and the necessary accessory components (multiplexer 16, memory 13, clock generator consisting of PLL circuit 11 and quartz 10, etc.) may be composed of discrete components. In order to be incorporated in a wristwatch, however, the functions must be integrated in as few components as possible, which may result in a single application specific circuit 30 in the extreme case. For example, signal processors of the TMS 320C5x series (manufacturer: Texas Instruments) may be used, in which multiplexer 16 is already contained, inter alia, and Flash RAMs of the type AM29LV800 (manufacturer: Amdahl) having a capacity of 8 MBit. Such a memory capacity and the application of the compression method for hearing sample data according to the invention as described hereinafter allow to attain an uninterrupted operation of the monitor for approx. 7 days.

[0022] In view of energy consumption, it is advantageous if the hearing sample processing unit, more particularly signal processor 9, is only periodically switched on. If e.g. one hearing sample per minute is taken, it is sufficient according to the processing method of the present invention to switch on the power supply of the signal processor for some seconds (less than 5, e.g. 4 seconds) only. For this purpose, the power supply receives an on-signal 25 from clock circuit 2 during whose presence the hearing sample processing unit is supplied with current. A further reduction of the energy consumption is obtained by the fact that flash memory 13 is only supplied with the current required for the storing process for a short time, 3 milliseconds at the end of each processed hearing sample recording being sufficient in the case of the above-suggested type. The signal 26 required therefor is generated by signal processor 9. The program controlling the signal processor is contained in a separate program memory which may be integrated in the signal processor itself, so that the hearing sample processing operation can also be performed while flash memory 13 is off.

[0023] Hereinafter, the method for the processing of the hearing samples is described. After the recording of the ambient noise (microphone 18) and its analog-digital conversion according to known principles (A/D converter unit 19), a splitting into e.g. six frequency bands is performed (Fig. 2) which is effected by a hierarchical arrangement of low passes 30 - 35. The required high pass associated to each low pass is realized by a subtraction 36 - 41 of the output signals 42 - 47 from the respective input signals 48 - 53 of the low passes, the subtraction being effected by an addition of the inverted output signals 42 - 47 of low passes 30 - 35.

[0024] Low pass filters 30 to 35 are realized by a 19-digit convolution:

$$y_j = \sum_{i=0}^{18} a_i \, x_{j-i} \qquad (1)$$

where

j :	time index
y_j :	output value of the low pass filtering at the time j;
x_j :	input value for low pass filtering at the time j;
a_i :	coefficient of the convolution sequence;
$a_0 \ldots a_{18}$:	[0.03, 0.0, -0.05, 0.0, 0.06, 0.0, -0.11, 0.0, 0.32, 0.50, 0.32, 0.0, -0.11, 0.0, 0.06, 0.0, -0.05, 0.0, 0.03]

[0025] In the course of the splitting into the frequency bands or band signals (54), a first data reduction is already effected in that only every second value out of each sequence of output values of the high and low pass filterings is transmitted to the following low resp. high pass stage or to outputs 54 by the switches 55. Overall, this already allows to obtain a reduction of the data volume to 1/8. With the division into six bands used in the example, this results in a slight overcompensation of the accompanying increase of the data volume by a factor six.

[0026] A criterion for the design of the filters is that one band may contain the contents of every other band in a clearly attenuated form at the most. A reduction to the half at least may be considered as clearly attenuated. Ideally, the bands only contain residual portions of directly adjacent bands, portions which are near or below the resolution of the digital numerical representation even. In the preferred digital realization, this aim is attained by low pass filtering (convolution) and subsequent subtraction of the filtered proportion from the input signal of the low pass filter.

[0027] The treatment of the band signals 54 resulting from the division into bands is identical in each band, Figs. 3 and 4 showing the processing of only one band 56 in a representative manner.

[0028] Input signal 56, which is identical to output signal 54, is first squared in that it is supplied to the two inputs of a multiplier 57 in parallel. Except a proportionality factor, this squaring corresponds to a calculation of the energy content of the proportion of the ambient noise which is represented by signal 56. Energy values 58 are subjected to a low pass filtering. This filtering is realized by means of a convolution over 48 values:

$$y_j^e = \sum_{i=0}^{47} b_i \, x_{j-i}^e \qquad (2)$$

where

j :	time index of the y^e and x^e values;
x_j^e :	energy value 58 at the time j;
y_j^e :	output signal of the low pass filter 59 at the time j;
b_i :	the coefficients of the convolution sequence, wherein $b_0 = b_1 = \ldots = b_{47} = 1.00$.

[0029] Of the output values of low pass filter 59, only every 48th value is forwarded to the following differentiation 61 by switch 60. Overall, here, a data reduction to 1/48 of the input data volume is obtained by the formation of a mean value.

[0030] In differentiator 61, each incoming value is delayed by a time unit in delay unit 62. Delay unit 62 may e.g. be a FIFO waiting queue having a length of 1.

[0031] In adder 63, the undelayed values are added to the inverted, delayed values, so that the values of the differences between two successive input values of the differentiator 61 are available at the output 64. The differences refer to a determined, constant and known time shift which is given by the time units, and consequently represent an approximation of the derivative with respect to time.

[0032] The energy difference values 64 are subjected to the normalized quantization. On one hand, according to Fig. 4, the absolute value of the energy difference values is formed in absolute value unit 65. These absolute values are supplied to a maximum value detector 66 at the output 67 of which the greater one of the values supplied to its inputs 68 appears. Since the output signal from output 67 is fed back to one of the two inputs 68 by a single-stage delay circuit 69, the maximum value of all values received by absolute value unit 65 is formed at output 67. The maximum values pass through another switch 70 which only transmits every 32nd value, i.e. a value which is the greatest within a hearing sample (the hearing sample duration used in this embodiment results in 32 energy difference values 64 per hearing sample in each frequency band).

[0033] In a reciprocal-computing and multiplication unit 71, the number 128 (= 2^7) is divided by the maximum value of the hearing sample and the result is supplied to an input 72 of a multiplicator 73. The other input of multiplicator 73 is then successively supplied with the energy difference values 64 among which the maximum value has been deter-

mined. For this purpose, the difference values 64 are temporarily stored in a FIFO buffer 75. The result of the multiplication in multiplicator 73, whose values are comprised between -128 and +127, is converted by converter 76 into integers in the range D from 0 to 255, corresponding to a byte having 8 bits. These numbers are used as addresses in a look-up table (LUT) 77 where a number in the range W = 0 to 15, i.e. a four-digit binary number, is associated to each input value. The discrete mapping of 8-bit numbers onto 4-bit numbers performed in LUT 77 is nonlinear and so designed that the resolution of small input numbers is finer than that of greater input values, i.e. that small input values are more emphasized. This may be referred to as a non-equidistant quantization.

[0034] The 4-bit values from output 78 are stored in flash memory 13 (Fig. 1).

[0035] The described normalized, non-equidistant quantization and compression unit is provided for each band according to the illustration of Fig. 3, resulting in 4-bit values for a total of 32 x 48 x 8 = 12,288 values per processing cycle which are recorded by the A/D converter at input 48 (Fig. 2). With an A/D conversion rate of 3,000 to 5,000 conversions per second, as provided by the currently available A/D converters of the lowest power consumption, this results in a hearing sample duration of approx. 2.5 to 4 s. With a supposed rate of one hearing sample per minute, the necessary memory capacity for the data amounts to 32 x 6 x 4 = 768 bit/min or 1'105'920 bit/d. The indicated 8 Mbit memory thus allows to record approx. 7 days of uninterrupted operation of the monitor.

[0036] In view of a reduction of the required computing, all cited calculations are effected by integer or fixed point arithmetic unless especially indicated, in particular an exponential representation of floating point numbers is avoided. The number of bits used for the representation of a number essentially depends on the used processor and on the data length provided by the latter. The above-mentioned processor family TMS320C5x uses 16-bit arithmetic. The binary point for fixed point arithmetic is set in such a manner that the limited computing accuracy is optimally utilized in each processing step although the probability of a data overflow is extremely low. Therefore, the binary point is set differently in the different processing steps. In the preferred embodiment of the band division, the least significant bit represents the value 2^{-16} for the filter coefficients and the value 2^{0} for the data values. Energy conversion and energy filtering are calculated by 32-bit integer arithmetic which is implemented as standard library function calls.

[0037] Prior to the storage in the flash memory or alternatively in the evaluating center, usual compression methods may be additionally applied which allow restoration of the original data in an identical form when decompressed.

[0038] In preparation of the recognition of the program elements which are possibly contained in the hearing samples, program samples are as exactly simultaneously as possible taken, e.g. directly at the broadcasting station, and stored. Prior to their comparison, the program samples are preferably subjected to the same processing and compression process as the hearing samples. This may be the case before the storage or only at the time of reading resp. playback of the stored program samples.

[0039] For the recognition, one of the usual correlation methods may be used. It is also possible to apply a coarse correlation using a fast computing procedure first and to perform a more precise and complicated correlation only if a sufficient probability of the presence of a given hearing sample has been found. In particular, such a preceding coarse correlation also provides a first coarse estimate of a subsisting minimal time shift between the hearing sample and the reference samples recorded at the station. In the more complex procedure, finer time shifts are analyzed and a more rugged comparison method is applied which takes account of the statistical distribution of the program signal and of interference signals.

[0040] Essentially, in the course of the evaluation, the simultaneous captured samples of each program as recorded each by a stationary unit are compared to the hearing samples of each monitor. An exemplary comparison method is illustrated in the following pseudocode which describes the correlation of a hearing sample of a monitor:

```
Decompress data of the monitor
OptimumMatch := -1

FOR StationaryUnit := 1 TO NumberOfStationaryUnits DO
        Load digitized program samples which have been recorded at the same
        time as the hearing samples of the monitor;
```

```
        FOR TimeShift := 1 TO MaxTimeShift STEP Timestep DO
        {Takes account of running inaccuracies of the timers by a step size of
        Timestep}
        Calculate matching coefficient c_t with standard correlation for the
        actual time shift and assign result to the variable ActualMatch;
                IF (ActualMatch > OptimumMatch) DO
                    OptimumMatch := ActualMatch;
                    OptimumTimeShift := TimeShift;
                    OptimumStationaryUnit := Stationary Unit;
                ENDIF
        ENDFOR
ENDFOR

IF(OptimumMatch > Threshold) DO
        RadioStation is recognized;
        The correct station is stored in the memory OptimumStationaryUnit
ELSE
        None of the surveyed reference programs was heard at this time
ENDIF
```

[0041] In this procedure, only one of the radio programs registered in 'NumberOfStationaryUnits' is determined in the hearing sample of a monitor, namely the one which yields the highest probability (value of the variable 'Optimum-Match').

[0042] In particular, the optional, unequivocally reversible compression of the hearing samples processed according to the invention is reversed. This is followed by the initialization of 'OptimumMatch' to the lowest value which also indicates "no match", i.e. the wearer of the monitor has listened to none of the monitored programs.

[0043] The program samples of each stationary unit simultaneously recorded with the current hearing sample (loop "For StationaryUnit:= 1 to NumberOfStationaryUnits ... EndDo" are StationaryUnit:= 1 to NumberOfStationaryUnits ... EndDo" are loaded and processed in the same manner as the hearing sample. Due to subsisting small time shifts between the hearing samples and the program samples, the following comparison is performed for a certain number 'MaxTimeShift' of assumed time shifts (loop "For TimeShift := 1 to MaxTimeShift ... Endfor"). The comparison is effected by a standard correlation of program and hearing sample data which are shifted forwards or backwards with respect to each other according to the 'TimeShift' variable. In order to always allow a full correlation over all values of the hearing sample, the program samples are therefore recorded over a longer period per sample, the beginning being additionally set earlier in time by the corresponding maximum time shift. Correspondingly, the length of the program sample is chosen in such a manner that the hearing sample is still completely contained in the program sample time even if the beginnings of the program sample and of the hearing sample are maximally displaced.

[0044] The normalized correlation is performed according to the following formula:

$$c_t = \frac{\sum_{i=1}^{N} (s_i \; m_{i-t})}{\sqrt{\sum_{i=1}^{N} (s_i)^2} \sqrt{\sum_{i=1}^{N} (m_{i-t})^2}} \qquad (3)$$

where

- t : time shift index (= 'TimeShift' in pseudocode);
- N : number of correlated values, generally equal to the number of values in a hearing sample;
- i : time index;
- S_i : hearing sample value at the time i;
- m_{i-t} : program sample value at the time i, displaced by t time steps;
- c_t : correlation value for the time shift t: $-1 \le c_t \le 1$.

[0045] The c_t values for different t values and program samples are compared, and the greatest c_t value overall is stored along with the indications of the conditions in which it has been recorded. These indications consist of the time shift, the stationary unit, i.e. the program, and of the correlation value c_t itself.

[0046] If the so determined greatest c_t value is superior to a predetermined threshold value, the corresponding program is considered to be contained in the hearing sample. If the threshold value is not attained, it is assumed that no one of the programs was heard.

[0047] Since the correlation must be performed correspondingly often due to the considerable scope of time shifts (t resp. TimeShift), a simplified alternative is conceivable where the time intervals are treated with a coarser graduation. For those c_t values which exceed a predetermined threshold, the correlation is repeated with a more rugged method while taking account of all detected time shifts.

[0048] A suitable rugged correlation is

$$r_t = \frac{\sum_{i=1}^{N} \left| s_i - a * m_{i-t} \right|}{\sum_{i=1}^{N} \left| s_i \right|} \qquad (4)$$

where

- r_t : "rugged" correlation value;
- a : scaling factor which takes account of the attenuation of the program signal with respect to the hearing sample;

the remaining symbols corresponding to formula (3).

[0049] The procedure thus essentially uses absolute values both of the deviation between the hearing sample and the scaled program signal and of the hearing sample signal. The scaling factor $\underline{a}$ is iteratively determined in such a manner that the rugged correlation value r_t becomes minimal. Compared to the normal correlation, large deviations are less weighted in the rugged correlation, thus taking account of statistical distributions of hearing sample values and of program signal values and therefore resulting in better recognition rates for real signals than the normal correlation value c_t. In particular, individual hearing samples with large deviations are less weighted.

[0050] Tests show that the described method not only eliminates or at least strongly reduces known interference effects such as secondary noise and time shifts but that damping (speakers, transmission lines, general acoustic conditions) and echo as well have only little influence on the recognition of a program. It has been particularly surprising to find that the program could often be detected in the hearing samples even when the program element was inaudible. The suppression of echo effects is attributed to the formation of a temporal mean (filter 59), in particular, especially if its time constant is chosen in such a manner as to be greater than the echo times usually found in a normal environment.

A typically frequency-dependent (acoustic) damping is compensated by the described suitable combination of a division into frequency bands, a normalization to the maximum value, and in taking into account of the damping by means of the scaling factor $\underline{a}$ in the calculation of r_t or by the calculation mode of c_t.

[0051] Modifications of the exemplary embodiment within the scope of the invention are apparent to those skilled in the art.

[0052] According to the technological development, different components (signal processors, memories, etc.) may be used. Alternatives are conceivable in particular for the flash memory, e.g. battery-backed up CMOS memories. The criteria, especially for portable monitors such as wristwatches, are an extended uninterrupted monitoring period and a minimal energy consumption. In certain circumstances it may be better to use a fast processing unit having a higher power dissipation if the higher energy consumption with respect to a slower unit is more than compensated by only temporary operation with intermediate inactive pauses. Besides the complete shut-off, many components such as e. g. the TMS320C5xx also offer special power saving modes. Also, the reduction of the clock rate of a fast unit often allows an important reduction of the energy consumption.

[0053] Depending on the used technology, different degrees of accuracy or numbers of digits of the binary numbers may be used. In tests, a sufficiently safe program recognition has been obtained with 4-bit end results. It is also conceivable, however, to effect a reduction to 3 bits, or to provide a greater number, e.g. 6 bits, 7 bits, or 8 bits. Greater numbers of binary digits are possible in particular if shorter wearing times are allowed or if memories of greater capacity become available.

[0054] In the case of higher numbers of digits of the end result, it may also be necessary to increase the number of digits in the preceding steps to the number of digits of the end result at least.

[0055] Mostly, the exact values for the nonlinear mapping by table 77 as well as the threshold values for the weighting of the correlation values can only be determined empirically. Although a function similar to a logarithmization is preferred, other functions are possible. It is also conversely conceivable to emphasize the greater values in D and to suppress the small values of the energy differences.

[0056] The factors and the number of digits of the convolutions may as well be chosen differently, and a different number of frequency bands into which the hearing samples are split is possible. In particular, it is conceivable in the case of modified A/D conversion speeds, different settings with respect to echo and/or damping compensation, or modified hearing sample durations, to adapt low pass 59, e.g. by changing the number of tabs of the convolution.

[0057] It is also conceivable to perform the analog-digital conversion at a later stage of the compression, particularly if the corresponding analog circuits offer advantages with respect to the processing speed or the space consumption in the monitor. In the extreme case, the digitization might be effected only immediately prior to the storage in the memory. If an analog signal is concerned, the term "digital value" in the description shall be replaced with e.g. the size or the amplitude of the signal.

[0058] With respect to the correlation, it is also possible to use only the part of the hearing samples which still lies within the corresponding program sample with the actual time shift t, e.g. if program and hearing samples of the same length are recorded.

[0059] An alternative of the wearing sensor consists of using currently available motion sensors. A known embodiment contains a contact which switches between the open and the closed state on motion but remains in one of the two states in the absence of motion.

Glossary

[0060]

Flash RAM	RAM (see there) which also conserves data in case of power failure but allows faster storage and easier erasure than classic non-volatile memories (PROM/EPROM).
RAM	read/write memory
time index	number of a digital value in the succession of values leaving the digitizer (A/D converter), mostly in relation to the beginning of a hearing sample, whose associated value has the time index 0.

Claims

1. Method for the compression of an electric audio signal representing hearing samples, wherein the amplitude of each sample as a whole or in parts, or of a digital or analog signal derived thereof is normalized to a first range D (65 - 76) of digital values, **characterized in that**

   - the hearing samples are created by recording of ambient noise by means of an electroacoustic transducer;

- said range D is predetermined;
- said normalized signal is mapped using a nonlinear function (77) onto a second predetermined range of digital values W yielding a result (78), the digital representation of the values of the range W comprising less digits than the digital representation of the values of the range D and the nonlinear function having a slope dW/dD which decreases with increasing values in order to obtain an emphasis of the small values of said first range of values; and
- the result (78) is stored in an electronic memory (13) in a digital form,

so that a reduction of the amount of data to be stored as the result is attained and the result still allows the recognition of program elements comprised in the hearing samples by comparison with program samples representing the program elements.

2. The method of claim 1, wherein said result (78) is represented by binary numbers having a fixed number of binary digits from 3 to 16 bits, preferably from 4 to 8 bits, and more preferably of 4 bits.

3. The method of one of claims 1 to 2, wherein said audio signal is divided into at least two band signals (56) by filtering (30 - 35, 36 - 41), each one of the band signals containing a frequency range of the audio signal, and each band signal only containing the content of the other band signals not at all or in a clearly attenuated form, more particularly attenuated to at most the half.

4. The method of claim 3, wherein 3 to 15, preferably 4 to 10, more preferably 5 to 8, and particularly preferably 6 band signals are produced.

5. The method of claim 3 or 4, wherein said band signals contain frequency ranges of the same width each, and all frequency ranges are comprised in the range of 500 Hz to 10,000 Hz.

6. The method of one of claims 3 to 5, wherein the band signals are generated by a single or a cascaded multiple splitting of an input signal (49 - 53), which is the audio signal (48) or one of the output signals (49 - 53), in applying the following steps:

- first low pass filtering (30 - 35) generating a first output band signal (49 - 47),
- subtraction (36 - 41) of the first output band signal from the input signal (48 - 53) for the generation of a second output band signal.

7. The method of claim 6, wherein all first low pass filterings (30 - 35) have the same Q-factor.

8. The method of one of claims 6 to 7, wherein said low pass filtering (30 - 35) is realized by means of a digital convolution over 10 - 30 values, preferably 15 - 25 values, and more preferably 19 values.

9. The method of claim 8, wherein the digital convolution is performed with the terms a_i*x_{t-i}, X_{t-i} being the input value for the convolution at a time preceding time t by i time periods, i being greater than or equal to 0 and less than or equal to 18, the coefficients a_i, being approximately equal to {0.03, 0.0, -0.05, 0.0, 0.06, 0.0, -0.11, 0.0, 0.32, 0.50, 0.32, 0.0, -0.11, 0.0, 0.06, 0.0, -0.05, 0.0, 0.03}.

10. The method of one of claims 6 to 9, wherein the input signal is digitized and only every nth value (55) of each division stage (30, 36; 31, 37; 32, 38; ...; 35, 41) is added to the band signal, n being at least 2 or equal to 2, in order to compensate for the increased data volume resulting from the splitting into band signals.

11. The method of one of claims 1 to 10, wherein an energy signal (58) which is proportional to the energy content is generated from said audio signal (48) or from a signal derived therefrom (54).

12. The method of claim 11, wherein the energy signal (58) is generated by squaring of the audio signal (48) or of a signal derived thereform (54).

13. The method of one of claims 11 to 12, wherein said energy signal (58) is subjected to a second low pass filtering.

14. The method of claim 13, wherein said second low pass filtering (59) is effected digitally in the form of a convolution over 20 to 70 values, preferably 40 to 55 values, and more preferably 48 values.

15. The method of claim 14, wherein the coefficients of the convolution are essentially equal to each other.

16. The method of claim 15, wherein the coefficients are about 1.0.

17. The method of one of claims 14 to 16, wherein said second low pass filtering is followed by a data reduction (60) where one energy value among n filtered values is selected, n being at least equal to 2 and preferably equal to the number of values of the convolution of the second low pass filtering (59).

18. The method of one of claims 11 to 17, wherein a subsequent differentiation of the energy signal with respect to the time (61) is effected in order to obtain an energy difference signal (64).

19. The method of claim 18, wherein said differentiation is effected by computing the difference between each two consecutive values of the signal.

20. The method of one of claims 1 to 19, wherein the normalization to a range of values D, which is defined by a lower limit D_u, preferably 0, and an upper limit D_o, is effected by:

- obtaining the maximum (67) of the absolute value (68) of the input signal within the normalizing duration of the signal, which is shorter or preferably equal to the duration of a hearing sample,
- by multiplying the reciprocal value of said maximum by (D_0 - D_u + 1) (71), and
- by multiplying this product by the values of the input signal (64) within the duration of the normalized signal.

21. The method of claim 20, wherein D_o - D_u is equal to 2^n-1, n being a whole number greater than 4 and preferably equal to 7.

22. The method of one of claims 1 to 21, wherein all steps of the method are performed by integer or fixed point arithmetic using a predetermined number of digits.

23. The method of claim 22, wherein the number of digits is the number of digits as provided by the employed computing unit (9).

24. Device (1) comprising a computer program product for carrying out the method of one of claims 1 to 23, wherein the device includes a hearing sample unit comprising at least one signal processor (9), the computer programm product comprising instructions for causing the signal processor to execute all the method steps of one of claims 1 to 23.

25. The device of claim 24, wherein a non-volatile semiconductor memory (13) is connected to said processor (9) which allows to store the results (78) of the method.

26. The device of claim 24 or 25, wherein a timer (2) is connected to the power supply (20) of said hearing sample unit which allows to switch off the hearing sample unit when no processing activity is required in the periods between the processing of two hearing samples, in order to reduce the energy consumption.

27. The device of claim 26, wherein the power supply of said non-volatile memory (13) and/or said memory itself is connected to a timer (2) in such a manner that the memory is capable of being operated only during the storage of the results in order to reduce the energy consumption by the memory.

28. An appliance normally worn by a person, **characterized in that** it comprises the device of one of claims 24 to 27, with the device being sufficiently small to be worn by a person.

29. The appliance of claim 28, **characterized in that** the appliance is a wristwatch.

30. Method for the evaluation of the results of registering hearing samples **comprising:**

carrying out the method of one of claims 1 to 23, **recording** program samples of monitored programs which **program samples** have at least the same duration as the hearing samples, subjecting the program samples to the same processing steps as the hearing samples, and **effecting** a calculation of a first correlation of the hearing samples with the processed program samples in order to find a match.

31. The method of claim 30, wherein the recording of the program samples is started before that of the hearing samples, its duration is longer than that of the hearing samples, and wherein in the correlation, time shifts between the timer for the hearing samples and the timer for the program samples are compensated by a displacement in time of the hearing samples with respect to the program samples.

32. The method of claim 30 or 31, wherein said first correlation is a standard correlation according to the formula

$$c_t = \frac{\sum_{i=1}^{N}\left(s_i\, m_{i-t}\right)}{\sqrt{\sum_{i=1}^{N}\left(s_i\right)^2}\sqrt{\sum_{i=1}^{N}\left(m_{i-t}\right)^2}}$$

where

N : number of values of the hearing sample which are used in the correlation,
t : time shift
s_i : hearing sample value at the time i,
m_i : program sample value at the time i,
c_t : correlation value for the time shift t: $-1 \leq c_t \leq 1$.

33. The method of one of claims 30 to 32, wherein the comparison of the hearing samples with the program samples is effected in two passes, a respective hearing sample being compared to all program samples in all ways in the first pass by means of said first correlation whose calculation demand is reduced by applying coarser graduation of time shifts by skipping time shift values, while in the case of a time shift whose correlation values c_t are above a predetermined limit, a second, rugged correlation is performed in skipping less time shift values, preferably no time shift value, whereby an improved graduation of the time shift is provided, in particular at least twice as high as in the first correlation.

34. The method of claim 33, wherein the second correlation is chosen such that great deviations between the hearing and the program sample have a smaller influence upon the correlation coefficients than in the first correlation.

35. The method of one of claims 33 to 34, wherein the second correlation is calculated according to the formula

$$r_t = \frac{\sum_{i=1}^{N}\left|s_i - a * m_{i-t}\right|}{\sum_{i=1}^{N}\left|s_i\right|}$$

where

N : number of hearing sample values used in the correlation,
t : time shift between the hearing and the program sample,
S_i : hearing sample value at the time i,
m_i : program sample value at the time i, and
a : scaling factor which takes account of the damping of the program signal with respect to the hearing sample;
r_t : correlation value for the shift t, 0 (optimal correlation) $\leq r_t \leq$ 1 (no correlation),

a being determined in such a manner that r_t assumes a minimal value.

36. Data carrier containing a recorded computer program product upon whose execution by a signal processor the method according to one of claims 1 to 23 and/or one of claims 30 to 35 is carried out.

Patentansprüche

1. Verfahren zur Kompression eines elektrischen Audiosignals, welches Hörproben darstellt, wobei die Amplitude der Proben als Ganzes oder in Teilen, bzw. eines davon abgeleiteten digitalen oder analogen Signals, jeweils auf einen ersten Bereich D (65 - 76) digitaler Werte normalisiert wird, **dadurch gekennzeichnet, dass**

   - die Hörproben durch Aufnehmen von Umgebungsgeräuschen mittels eines elektroakustischen Wandlers erzeugt werden;
   - der genannte Bereich D vorgegeben ist;
   - das genannte normalisierte Signal unter Verwendung einer nichtlinearen Funktion (77) auf einen zweiten festgelegten Bereich digitaler Werte W abgebildet wird, der ein Resultat (78) liefert, wobei die digitale Darstellung der Werte im Bereich W weniger Stellen enthält als die digitale Darstellung der Werte im Bereich D und die nichtlineare Funktion eine Steigung dW/dD aufweist, welche mit steigenden Werten abnimmt, um kleine Werte des genannten ersten Wertebereichs hervorzuheben; und
   - das Resultat (78) in digitaler Form in einem elektronischen Speicher (13) gespeichert wird,

   so dass eine Reduktion der als Resultat zu speichernden Datenmenge erzielt wird und das Resultat immer noch die Erkennung von in den Hörproben enthaltenen Programmteilen gestattet, wenn diese mit Programmproben verglichen werden, welche die Programmteile darstellen.

2. Verfahren nach Anspruch 1, worin das genannte Resultat (78) durch binäre Zahlen mit einer festen Anzahl binärer Stellen von 3 bis 16 Bit dargestellt wird, bevorzugt von 4 bis 8 Bit, und weiter bevorzugt von 4 Bit.

3. Verfahren nach einem der Ansprüche 1 bis 2, worin das genannte Audiosignal durch Filtern (30 - 35, 36 - 41) in mindestens zwei Bandsignale (56) aufgetrennt wird, welche jeweils einen Frequenzbereich des Audiosignals enthalten, und jedes Bandsignal den Inhalt der anderen Bandsignale gar nicht oder nur in deutlich abgeschwächter Form enthält, insbesondere auf höchstens die Hälfte abgeschwächt.

4. Verfahren nach Anspruch 3, worin 3 bis 15, bevorzugt 4 bis 10, weiter bevorzugt 5 bis 8 und besonders bevorzugt 6 Bandsignale erzeugt werden.

5. Verfahren nach Anspruch 3 oder 4, worin die genannten Bandsignale jeweils Frequenzbereiche gleicher Breite enthalten und alle Frequenzbereiche im Bereich von 500 Hz bis 10'000 Hz liegen.

6. Verfahren nach einem der Ansprüche 3 bis 5, worin die Bandsignale durch eine einfache oder kaskadierte mehrfache Auftrennung eines Eingangssignals (49 - 53), bei welchem es sich um das Audiosignal (48) oder um eines der Ausgangssignale (49 - 53) handelt, unter Anwendung folgender Schritte erzeugt werden:

   - erste Tiefpassfilterung (30 - 35) zur Erzeugung eines ersten Ausgangs-Bandsignals (49 - 47),
   - Subtraktion (36 - 41) des ersten Ausgangs-Bandsignals vom Eingangssignal (48 - 53) zur Erzeugung eines zweiten Ausgangs-Bandsignals.

7. Verfahren nach Anspruch 6, worin alle ersten Tiefpassfilterungen (30 - 35) den gleichen Gütefaktor aufweisen.

8. Verfahren nach einem der Ansprüche 6 bis 7, worin die Tiefpassfilterung (30 - 35) durch eine digitale Faltung über 10 bis 30 Werte, vorzugsweise 15 bis 25 Werte und weiter bevorzugt über 19 Werte erfolgt.

9. Verfahren nach Anspruch 8, worin die digitale Faltung gemäss dem Ausdruck $a_i{}^*x_{t-i}$ erfolgt, wobei X_{t-i} der Eingangswert der Faltung zu einem i Zeitperioden vor dem Zeitpunkt t liegenden Zeitpunkt ist und i grösser oder gleich 0 und kleiner oder gleich 18 ist, und wobei die Koeffizienten a_i ungefähr {0,03, 0,0, -0,05, 0,0, 0,06, 0,0, -0,11, 0,0, 0,32, 0,50, 0,32, 0,0, -0,11, 0,0, 0,06, 0,0, -0,05, 0,0, 0,03} betragen.

10. Verfahren nach einem der Ansprüche 6 bis 9, worin das Eingangssignal digitalisiert wird und nur jeder n-te Wert (55) jeder Teilungsstufe (30, 36; 31, 37; 32, 38; ...; 35, 41) zum Bandsignal zugefügt wird, wobei n mindestens 2

oder gleich 2 ist, um die Vergrösserung des Datenvolumens durch die Auftrennung in Bandsignale zu kompensieren.

11. Verfahren nach einem der Ansprüche 1 bis 10, worin aus dem Audiosignal (48) oder aus einem davon abgeleiteten Signal (54) ein zum Energiegehalt proportionales Energiesignal (58) gewonnen wird.

12. Verfahren nach Anspruch 11, worin das Energiesignal (58) durch Quadrieren des Audiosignals (48) oder eines davon abgeleiteten Signals (54) erzeugt wird.

13. Verfahren nach einem der Ansprüche 11 bis 12, worin das Energiesignal (58) einer zweiten Tiefpassfilterung unterzogen wird.

14. Verfahren nach Anspruch 13, worin die zweite Tiefpassfilterung (59) digital in Form einer Faltung über 20 bis 70 Werte erfolgt, bevorzugt über 40 bis 55 Werte, weiter bevorzugt über 48 Werte.

15. Verfahren nach Anspruch 14, worin die Koeffizienten der Faltung jeweils im wesentlichen untereinander gleich sind.

16. Verfahren nach Anspruch 15, worin die Koeffizienten ungefähr 1.0 betragen.

17. Verfahren nach einem der Ansprüche 14 bis 16, worin auf die zweite Tiefpassfilterung eine Datenreduktion (60) folgt, bei welcher ein Energiewert unter n gefilterten Werten ausgewählt wird, wobei n mindestens gleich 2 und bevorzugt gleich der Anzahl Werte der Faltung der zweiten Tiefpassfilterung (59) ist.

18. Verfahren nach einem der Ansprüche 11 bis 17, worin eine nachfolgende Differenzierung des Energiesignals nach der Zeit (61) durchgeführt wird, um ein Energiedifferentialsignal (64) zu erhalten.

19. Verfahren nach Anspruch 18, worin die genannte Differenzierung durch Berechnung der Differenz zwischen je zwei aufeinanderfolgenden Werten des Signals erfolgt.

20. Verfahren nach einem der Ansprüche 1 bis 19, worin die Normalisierung auf einen Wertebereich D, der definiert ist durch eine Untergrenze D_u, vorzugsweise 0, und eine Obergrenze D_o, erfolgt, indem:

- das Maximum (67) des Absolutwerts (68) des Eingangssignals innerhalb der Normalisierungsdauer des Signals ermittelt wird, welche kürzer ist als die Dauer einer Hörprobe oder bevorzugt gleich,
- der Kehrwert des Maximums mit $(D_o - D_u + 1)$ (71) multipliziert wird und
- dieses Produkt mit den Werten des Eingangssignals (64) innerhalb der Dauer des normalisierten Signals multipliziert wird.

21. Verfahren nach Anspruch 20, worin $D_o - D_u$ gleich 2^n-1 ist, wobei n eine ganze Zahl grösser als 4 und bevorzugt gleich 7 ist.

22. Verfahren nach einem der Ansprüche 1 bis 21, worin alle Verfahrensschritte mittels Ganzzahl- oder Festpunktarithmetik mit einer festgelegten Anzahl Stellen ausgeführt werden.

23. Verfahren nach Anspruch 22, worin die Anzahl Stellen die in der verwendeten Recheneinheit (9) zur Verfügung stehende Anzahl Stellen ist.

24. Vorrichtung (1) mit einem Computerprogrammprodukt zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 23, worin die Vorrichtung eine Hörprobeneinheit mit mindestens einem Signalprozessor (9) aufweist, wobei das Computerprogrammprodukt Instruktionen beinhaltet, welche die Ausführung aller Verfahrensschritte nach einem der Ansprüche 1 bis 7 durch den Signalprozessor auslösen.

25. Vorrichtung nach Anspruch 24, worin ein nichtflüchtiger Halbleiterspeicher (13) am genannten Prozessor (9) angeschlossen ist, welcher die Resultate (78) des Verfahrens zu speichern gestattet.

26. Vorrichtung nach Anspruch 24 oder 25, worin an der Stromversorgung (20) der Hörprobeneinheit ein Zeitgeber (2) angeschlossen ist, welcher die Hörprobeneinheit abzuschalten gestattet, wenn zwischen der Verarbeitung zweier Hörproben keine Verarbeitungsaktivität erforderlich ist, um den Energieverbrauch zu verringern.

27. Vorrichtung nach Anspruch 26, worin die Stromversorgung des nichtflüchtigen Speichers (13) und/oder der Speicher selbst an einem Zeitgeber (2) angeschlossen ist, so dass ein Betrieb des Speichers nur während der Speicherung der Resultate ermöglicht ist, um den Energieverbrauch durch den Speicher zu verringern.

28. Gerät, welches normalerweise von einer Person getragen wird, **dadurch gekennzeichnet, dass** es eine Vorrichtung nach einem der Ansprüche 24 bis 27 aufweist, welche genügend klein ist, um von einer Person getragen zu werden.

29. Gerät nach Anspruch 28, **dadurch gekennzeichnet, dass** das Gerät eine Armbanduhr ist.

30. Verfahren zur Auswertung der Resultate aus der Erfassung von Hörproben, beinhaltend: die Durchführung des Verfahrens nach einem der Ansprüche 1 bis 23, die Aufzeichnung von Programmproben überwachter Programme, welche Programmproben mindestens die gleiche Zeitdauer aufweisen wie die Hörproben, die Anwendung der gleichen Bearbeitungsschritte auf die Programmproben wie auf die Hörproben, und die Durchführung der Berechnung einer ersten Korrelation der Hörproben mit den bearbeiteten Programmproben, um eine Übereinstimmung zu finden.

31. Verfahren nach Anspruch 30, worin die Aufzeichnung der Programmproben vor derjenigen der Hörproben beginnt und länger dauert als diejenige der Hörproben, und worin Zeitverschiebungen zwischen dem Zeitgeber für die Hörproben und dem Zeitgeber für die Programmproben bei der Korrelation durch zeitliche Verschiebung der Hörproben gegenüber den Programmproben kompensiert werden.

32. Verfahren nach Anspruch 30 oder 31, worin die genannte erste Korrelation eine Standardkorrelation nach der Formel

$$c_t = \frac{\sum_{i=1}^{N} (s_i \, m_{i-t})}{\sqrt{\sum_{i=1}^{N} (s_i)^2} \sqrt{\sum_{i=1}^{N} (m_{i-t})^2}}$$

ist, mit

- N : Anzahl Werte der Hörprobe, welche in der Korrelation verwendet werden,
- t : Zeitverschiebung
- s_i : Wert der Hörprobe zum Zeitpunkt i,
- m_i : Wert der Programmprobe zum Zeitpunkt i,
- c_t : Korrelationswert für die Zeitverschiebung t: $-1 \le c_t \le 1$.

33. Verfahren nach einem der Ansprüche 30 bis 32, worin der Vergleich der Hörproben mit den Programmproben in zwei Durchgängen erfolgt, wobei die Hörproben im ersten Durchgang jeweils auf alle Arten mit allen Programmproben verglichen werden mit Hilfe der genannten ersten Korrelation, deren Rechenaufwand dadurch verringert wird, dass eine gröbere Abstufung der Zeitverschiebungen angewendet wird, indem Zeitverschiebungswerte übersprungen werden, während im Fall einer Zeitverschiebung, deren Korrelationswerte c_t über einer festgelegten Grenze liegen, eine zweite, stabilere Korrelation erfolgt, indem weniger oder bevorzugt keine Zeitverschiebungswerte übersprungen werden, wodurch eine verbesserte Abstufung der Zeitverschiebung erhalten wird, insbesondere eine dmindestens oppelt so feine Abstufung als in der ersten Korrelation.

34. Verfahren nach Anspruch 33, worin die zweite Korrelation derart gewählt ist, dass grosse Abweichungen zwischen der Hörprobe und der Programmprobe einen kleineren Einfluss auf die Korrelationskoeffizienten haben als in der ersten Korrelation.

35. Verfahren nach einem der Ansprüche 33 bis 34, worin die zweite Korrelation nach der Formel

$$r_t = \frac{\sum_{i=1}^{N} |s_i - a * m_{i-t}|}{\sum_{i=1}^{N} |s_i|}$$

berechnet wird, worin

N : Anzahl Hörprobenwerte, welche in der Korrelation verwendet werden,
t : Zeitverschiebung zwischen der Hörprobe und der Programmprobe,
s_i : Wert der Hörprobe zum Zeitpunkt i,
m_i : Wert der Programmprobe zum Zeitpunkt i, und
a : Skalierfaktor, der die Dämpfung des Programmsignals gegenüber der Hörprobe berücksichtigt;
r_t : Korrelationswert für die Zeitverschiebung t, 0 (optimale Korrelation) $\leq r_t \leq 1$ (keine Korrelation),

wobei a derart festgelegt wird, dass r_t einen Minimalwert annimmt.

36. Datenträger mit einem aufgezeichneten Computerprogrammprodukt, bei dessen Ausführung durch einen Signalprozessor das Verfahren nach einem der Ansprüche 1 bis 23 und/oder einem der Ansprüche 30 bis 35 durchgeführt wird.

Revendications

1. Procédé de compression d'un signal audio électrique représentant des échantillons d'écoute, où l'amplitude de chaque échantillon, en tout ou partie, ou d'un signal numérique ou analogique dérivé de ce dernier, est normalisée à un premier domaine D (65 - 76) de valeurs numériques, **caractérisé en ce que**

- les échantillons d'écoute sont créés en enregistrant des bruits d'environnement au moyen d'un traducteur électroacoustique;
- ledit domaine D est prédéterminé;
- ledit signal normalisé est appliqué à l'aide d'une fonction non linéaire (77) sur un deuxième domaine prédéterminé de valeurs numériques W fournissant un résultat (78), la représentation numérique des valeurs du domaine W comprenant moins de chiffres que la représentation numérique des valeurs du domaine D, et la fonction non linéaire présentant une pente dW/dD qui décroît avec des valeurs croissantes afin d'obtenir une accentuation de petites valeurs dudit premier domaine de valeurs; et
- le résultat (78) est mémorisé sous forme numérique dans une mémoire électronique (13),

de manière qu'une réduction de la quantité de données à mémoriser comme résultat est atteinte, et que le résultat permet néanmoins la reconnaissance d'éléments d'un programme compris dans les échantillons d'écoute par comparaison avec des échantillons de programme représentant lesdits éléments d'un programme.

2. Procédé selon la revendication 1, où ledit résultat (78) est représenté par des nombres binaires ayant un nombre fixe de chiffres binaires de 3 à 16 bits, préférablement de 4 à 8 bits, et plus préférablement de 4 bits.

3. Procédé selon l'une des revendications 1 à 2, où ledit signal audio est divisé en deux signaux de bande (56) au moins par filtrage (30 - 35, 36 - 41), chacun des signaux de bande contenant une gamme de fréquences du signal audio, et chaque signal de bande ne contenant pas du tout le signal des autres signaux de bande ou seulement sous forme clairement atténuée, plus particulièrement atténuée de plus de la moitié.

4. Procédé selon la revendication 3, où 3 à 15, préférablement 4 à 10, plus préférablement 5 à 8, et en particulier de préférence 6 signaux de bande sont produits.

5. Procédé selon la revendication 3 ou 4, où lesdits signaux de bande contiennent des gammes de fréquences de la même largeur chacun, et toutes les gammes de fréquences sont comprises dans la gamme de 500 Hz à 10'000

Hz.

6. Procédé selon l'une des revendications 3 à 5, où les signaux de bande sont générés par une division simple ou multiple en cascade d'un signal d'entrée (49 - 53), qui est le signal audio (48) ou l'un des signaux de sortie (49-53), en appliquant les étapes suivantes:

   - premier filtrage passe-bas (30 - 35) générant un premier signal de bande de sortie (49 - 47),
   - soustraction (36 - 41) du premier signal de bande de sortie du signal d'entrée (48 - 53) pour la génération d'un deuxième signal de bande de sortie.

7. Procédé selon la revendication 6, où tous les premiers filtrages passe-bas (30 - 35) ont le même facteur d'amplification dynamique Q.

8. Procédé selon l'une des revendications 6 à 7, où ledit filtrage passe-bas (30 - 35) est réalisé au moyen d'une convolution numérique sur 10 à 30 valeurs, préférablement 15 à 25 valeurs, et plus préférablement 19 valeurs.

9. Procédé selon la revendication 8, où la convolution numérique est effectuée selon le terme $a_i{}^*x_{t-i}$, X_{t-i} étant la valeur d'entrée de la convolution à un temps précédant le temps t de i périodes, i étant supérieur ou égal à 0 et inférieur ou égal à 18, les coefficients a_i étant approximativement égaux à {0,03, 0,0, -0,05, 0,0, 0,06, 0,0, -0,11, 0,0, 0,32, 0,50, 0,32, 0,0, -0,11, 0,0, 0,06, 0,0, -0,05, 0,0, 0,03}.

10. Procédé selon l'une des revendications 6 à 9, où le signal d'entrée est numérisé et seulement chaque n-ième valeur (55) de chaque étape de division (30, 36; 31, 37; 32, 38; ...; 35, 41) est ajoutée au signal de bande, n étant au moins 2 ou égal à 2, afin de compenser le volume de données accru en raison de la division en signaux de bande.

11. Procédé selon l'une des revendications 1 à 10, où un signal d'énergie (58) proportionnel au contenu énergétique est généré à partir dudit signal audio (48) ou à partir d'un signal (54) dérivé de ce dernier.

12. Procédé selon la revendication 11, où le signal d'énergie (58) est généré en formant le carré du signal audio (48) ou d'un signal (54) dérivé de ce dernier.

13. Procédé selon l'une des revendications 11 à 12, où ledit signal d'énergie (58) est soumis à un deuxième filtrage passe-bas.

14. Procédé selon la revendication 13, où ledit filtrage passe-bas (59) est effectué de manière numérique sous forme d'une convolution sur 20 à 70 valeurs, préférablement 40 à 55 valeurs, et plus préférablement 48 valeurs.

15. Procédé selon la revendication 14, où les coefficients de la convolution sont essentiellement égaux l'un à l'autre.

16. Procédé selon la revendication 15, où les coefficients sont égaux à environ 1,0.

17. Procédé selon l'une des revendications 14 à 16, où ledit deuxième filtrage passe-bas est suivi d'une réduction des données (60) dans laquelle une valeur énergétique parmi n valeurs filtrées est choisie, n étant égal à 2 au moins et préférablement égal au nombre de valeurs de la convolution du deuxième filtrage passe-bas (59).

18. Procédé selon l'une des revendications 11 à 17, où une différentiation du signal énergétique par rapport au temps (61) est effectuée afin d'obtenir un signal différentiel énergétique (64).

19. Procédé selon la revendication 18, où ladite différentiation est effectuée en calculant chaque fois la différence entre deux valeurs consécutives du signal.

20. Procédé selon l'une des revendications 1 à 19, où la normalisation à un domaine de valeurs D, qui est défini par une limite inférieure D_u, préférablement 0, et une limite supérieure D_o, est effectuée:

   - en obtenant le maximum (67) de la valeur absolue (68) du signal d'entrée au cours de la durée de normalisation du signal, qui est inférieure ou préférablement égale à la durée de l'échantillon d'écoute,
   - en multipliant la valeur réciproque dudit maximum par (D_o - D_u + 1) (71), et
   - en multipliant ce produit par les valeurs du signal d'entrée (64) pendant la durée du signal normalisé.

21. Procédé selon la revendication 20, où D_o - D_u est égal à 2^n-1, n étant un nombre entier supérieur à 4 et préférablement égal à 7.

22. Procédé selon l'une des revendications 1 à 21, où toutes les étapes du procédé sont effectuées par des procédures d'arithmétique en nombres entiers ou en virgule fixe utilisant un nombre prédéterminé de chiffres.

23. Procédé selon la revendication 23, où le nombre de chiffres est le nombre de chiffres disponible dans l'unité de calcul (9) utilisée.

24. Dispositif (1) comprenant un produit de logiciel d'ordinateur pour la mise en oeuvre du procédé selon l'une des revendications 1 à 23, où le dispositif comporte une unité à échantillons d'écoute comprenant un processeur de signaux (9) au moins, le produit de logiciel d'ordinateur comprenant des instructions occasionnant l'exécution de toutes les étapes du procédé selon l'une des revendications 1 à 7 par le processeur de signaux.

25. Dispositif selon la revendication 24, où une mémoire (13) non volatile à semi-conducteurs est connectée audit processeur (9), ladite mémoire permettant de mémoriser les résultats (78) du procédé.

26. Dispositif selon la revendication 24 ou 25, où une horloge (2) est connectée à l'alimentation (20) de ladite unité à échantillons d'écoute, ladite horloge permettant de désactiver l'unité à échantillons d'écoute lorsque aucune activité de traitement n'est requise dans les périodes entre le traitement de deux échantillons d'écoute, afin de réduire la consommation d'énergie.

27. Dispositif selon la revendication 26, où l'alimentation de ladite mémoire non volatile (13) et/ou ladite mémoire elle-même est connectée à une horloge (2) de telle manière que la mémoire est capable d'être enclenchée pendant la mise en mémoire des résultats uniquement, afin de réduire la consommation d'énergie de la mémoire.

28. Appareil normalement porté par une personne, **caractérisé en ce qu'**il comporte le dispositif selon l'une des revendications 24 à 27, ledit dispositif étant suffisamment petit pour être porté par une personne.

29. Appareil selon la revendication 28, **caractérisé en ce que** l'appareil est une montre-bracelet.

30. Procédé d'évaluation des résultats de la saisie d'échantillons d'écoute, comprenant: la mise en oeuvre du procédé selon l'une des revendications 1 à 23, l'enregistrement d'échantillons de programme de programmes surveillés, les échantillons de programme ayant la même durée au moins que les échantillons d'écoute, l'application aux échantillons de programme des mêmes étapes de traitement qu'aux échantillons d'écoute, et la mise en oeuvre d'un calcul d'une première corrélation des échantillons d'écoute avec les échantillons de programme traités afin de trouver une concordance.

31. Procédé selon la revendication 30, où l'enregistrement des échantillons de programme commence avant celui des échantillons d'écoute, sa durée est plus longue que celle des échantillons d'écoute, et où dans la corrélation, des décalages temporels entre l'horloge pour les échantillons d'écoute et l'horloge pour les échantillons de programme sont compensés par un déplacement dans le temps des échantillons d'écoute par rapport aux échantillons de programme.

32. Procédé selon la revendication 30 ou 31, où ladite première corrélation est une corrélation standard selon la formule

$$c_t = \frac{\sum_{i=1}^{N} (s_i \, m_{i-t})}{\sqrt{\sum_{i=1}^{N} (s_i)^2} \sqrt{\sum_{i=1}^{N} (m_{i-t})^2}}$$

où

N : nombre de valeurs de l'échantillon d'écoute utilisées dans la corrélation,
t : décalage temporel,
S_i : valeur de l'échantillon d'écoute au moment i,
m_i : valeur de l'échantillon de programme au moment i,
c_t : valeur de la corrélation pour le décalage temporel t: $-1 \le c_t \le 1$.

33. Procédé selon l'une des revendications 30 à 32, où la comparaison des échantillons d'écoute aux échantillons de programme est effectuée en deux passages, un échantillon d'écoute donné étant comparé à tous les échantillons de programme, dans tous les cas, lors du premier passage au moyen de ladite première corrélation, dont les exigences arithmétiques sont réduites en sautant des valeurs de décalages temporels, tandis que dans le cas d'un décalage temporel dont les valeurs de corrélation c_t sont supérieures à une limite prédéterminée, une deuxième corrélation plus robuste est effectuée en sautant moins de valeurs de décalages temporels, préférablement aucune valeur de décalage temporel, fournissant ainsi une meilleure gradation des décalages temporels, plus particulièrement une gradation au moins deux fois plus fine que dans la première corrélation.

34. Procédé selon la revendication 33, où la deuxième corrélation est choisie telle que des grandes déviations entre l'échantillon d'écoute et l'échantillon de programme ont une moindre influence sur les coefficients de corrélation que dans la première corrélation.

35. Procédé selon l'une des revendications 33 à 34, où la deuxième corrélation est calculée selon la formule

$$r_t = \frac{\sum_{i=1}^{N} |s_i - a * m_{i-t}|}{\sum_{i=1}^{N} |s_i|}$$

où

N : nombre de valeurs de l'échantillon d'écoute utilisées dans la corrélation,
t : décalage temporel entre l'échantillon d'écoute et l'échantillon de programme,
s_i : valeur de l'échantillon d'écoute au moment i,
m_i : valeur de l'échantillon de programme au moment i, et
a : facteur de cadrage tenant compte de l'atténuation du signal de programme par rapport à l'échantillon d'écoute;
r_t : valeur de la corrélation pour le décalage t, 0 (corrélation optimale) $\le r_t \le$ 1 (aucune corrélation),

a étant déterminé de telle manière que r_t prend une valeur minimale.

36. Support de données contenant un produit de logiciel d'ordinateur dont l'exécution par un processeur de signal met en oeuvre le procédé selon l'une des revendications 1 à 23 et/ou l'une des revendications 30 à 35.

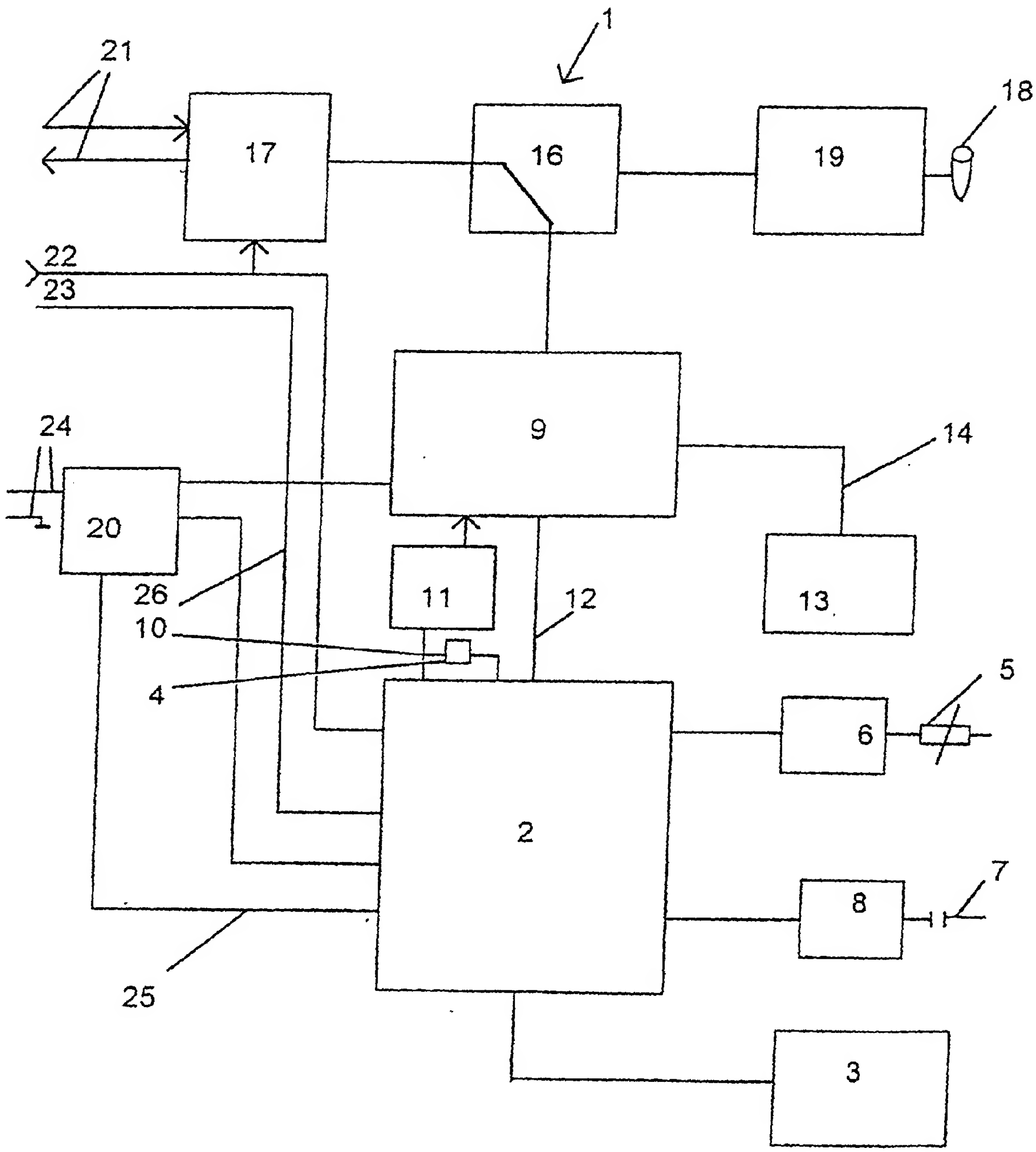

Fig. 1

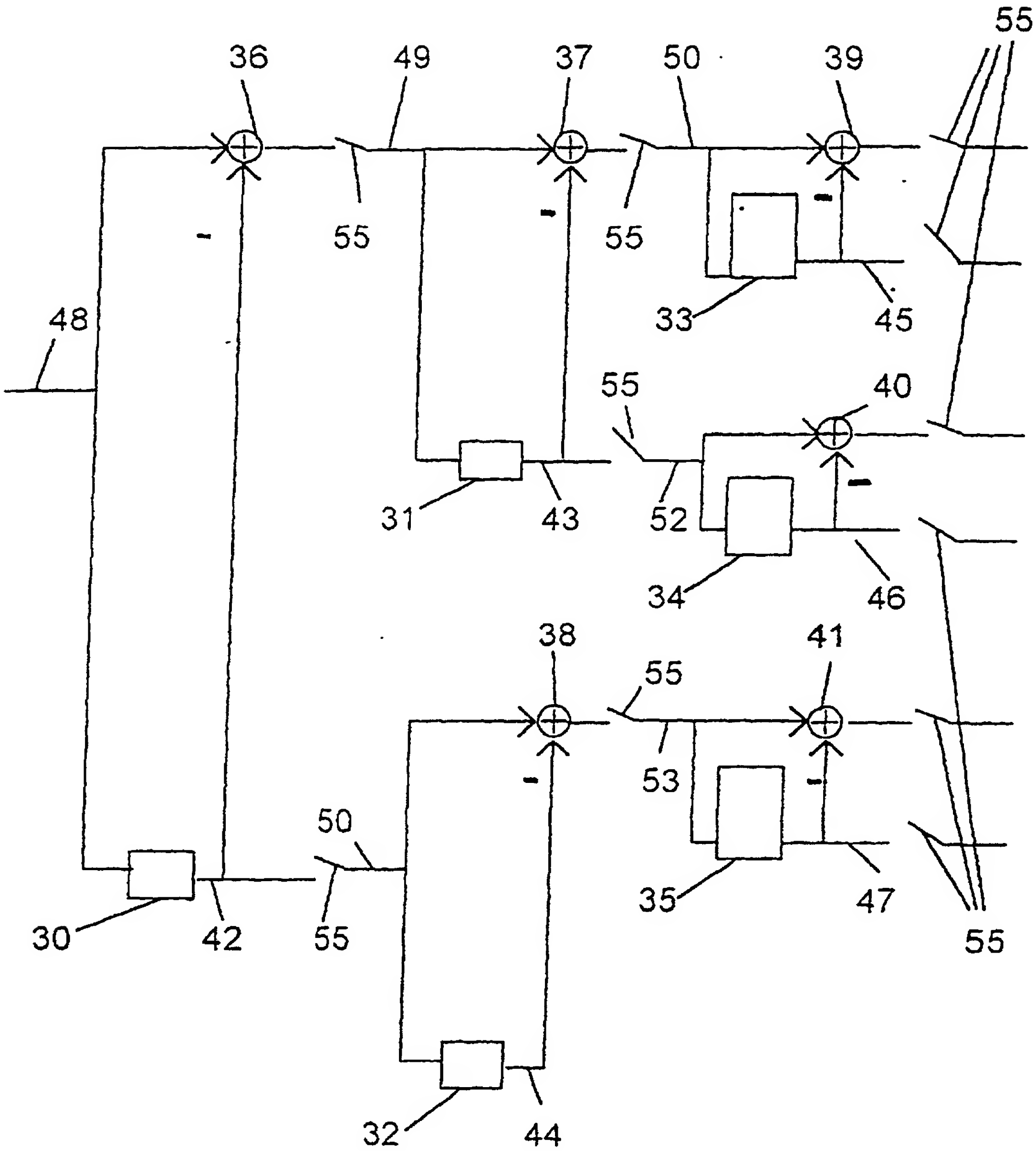
55
36
49
37
50
39
55
55
48
33
45
55
40
31
43
52
34
46
38
41
55
53
50
35
47
30
42
55
55
32
44

Fig. 2

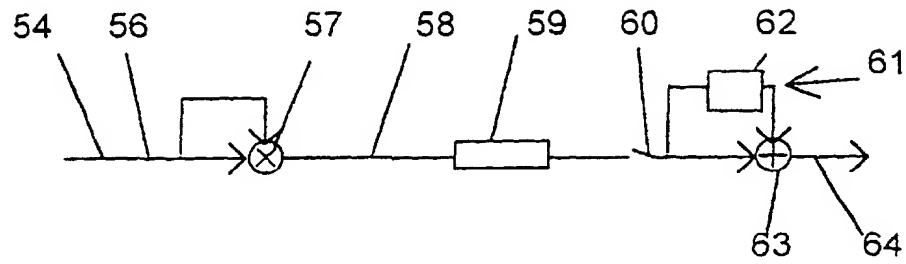
54
56
57
58
59
60
62
61
63
64

Fig. 3

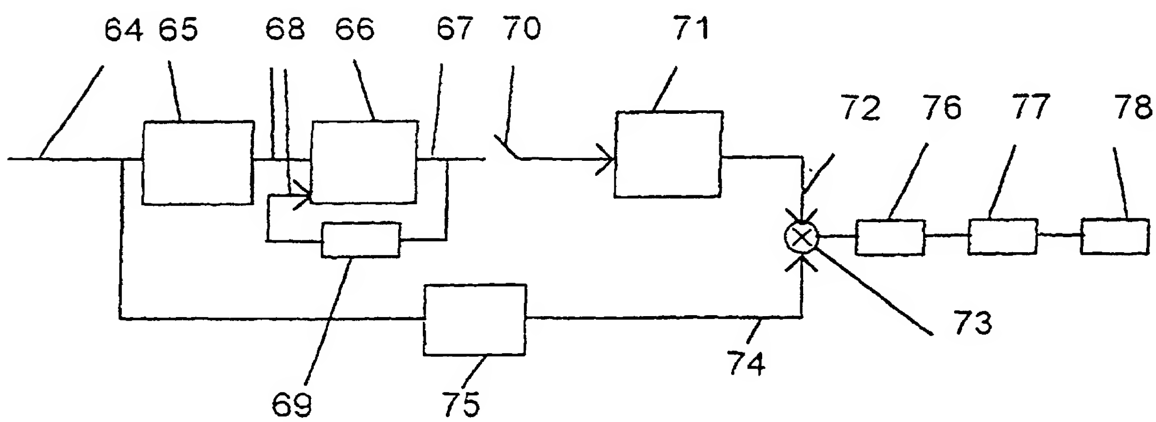
64
65
68
66
67
70
71
72
76
77
78
73
74
69
75

Fig. 4